haematologica
Journal of the Ferrata Storti Foundation

Evaluation of acute myeloid leukemia using genomic proximity mapping-based next generation cytogenomics

by Cecilia C.S. Yeung, Stephen M. Eacker, Olga Sala-Torra, Mary Wood, Lan Beppu, David W. Woolston, Ivan Liachko, Maika Malig, Derek Stirewalt, Alexander Muratov, Min Fang and Jerald Radich

Received: June 17, 2025.
Accepted: January 7, 2026.

Citation: Cecilia C.S. Yeung, Stephen M. Eacker, Olga Sala-Torra, Mary Wood, Lan Beppu, David W. Woolston, Ivan Liachko, Maika Malig, Derek Stirewalt, Alexander Muratov, Min Fang and Jerald Radich. Evaluation of acute myeloid leukemia using genomic proximity mapping-based next generation cytogenomics. Haematologica. 2026 Jan 15. doi: 10.3324/haematol.2025.288461 [Epub ahead of print]

Evaluation of acute myeloid leukemia using genomic proximity mapping-based next generation cytogenomics

Cecilia CS Yeung[1,2]*, Stephen M. Eacker[3]*, Olga Sala-Torra[1], Mary Wood[3], Lan Beppu[1], David W. Woolston[1], Ivan Liachko[3], Maika Malig[3], Derek Stirewalt[1,4], Alexander Muratov[3], Min Fang[1,2]**, Jerald Radich[1,4]**

*Co-First

**Co-Last

Affiliations:

1. Translational Science and Therapeutics Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

2. Department of Laboratory Medicine and Pathology, University of Washington Medical Center, Seattle, WA, USA

3. Phase Genomics, Seattle, WA, USA

4. Department of Medicine, University of Washington Medical Center, Seattle, WA, USA

Running head: Evaluation of AML using genomic proximity mapping.

Corresponding author:

Cecilia CS Yeung, M.D., FCAP
Fred Hutchinson Cancer Center
1100 Fairview Avenue North, G7-910, Seattle, WA 98109
Office Phone: (206) 606-1355
Email: cyeung@fredhutch.org
ORCID ID: 0000-0001-6799-2022

Data Sharing Statement

This is an active project with ongoing enrollment of subjects for the study thus the data will be made publicly available at a later date when the study is ended. However, data presented in this study will be made available upon request to the corresponding author, please contact cyeung@fredhutch.org.

Word count: 3567 words

Number of tables: 2

Number of figures: 6

Number of supplementary files: 7

Funding

This work was supported by an SBIR Phase II grant from NCI/NIH R44CA278140 to SME and Phase Genomics. This work is also partially supported by UG1 CA233338-02, and CA175008-06 to JR at FHCC.

Key words: Acute myeloid leukemia, AML, Hi-C sequencing, Genomic Proximity Mapping, Molecular genetics pathology, Genomic profiling

Acknowledgements

The authors thank the Genomics Core at Fred Hutchinson Cancer Center, especially Andy Marty and Alex Zevin, Ph.D., for the preparation and analysis of the WGS Illumina libraries.

Conflict of Interest Disclosure

IL, MM, and SME are employees of Phase Genomics, Inc., a company developing the GPM technology. CY consulted for TwinStrand Bioscience; this relationship has ended.

Author contributions

For every author, his or her contribution to the manuscript needs to be provided using the following categories:

Conception: CCSY, SME, IL, MF, JR, OST

Sample curation: OST, LB, DLS

Sample preparation and testing: LB, MM

Interpretation or analysis of data: CCSY, SME, OST, MW, DWW, IL, MM, DLS, AM, MF, JR

Preparation of the manuscript: CCSY, SME, OST, MW, LB, DWW, IL, MM, DLS, AM, MF, JR

Revision for important intellectual content: MF, JR

Supervision: CCSY, SME

Abstract

Cytogenetic analysis encompasses a suite of standard-of-care diagnostic testing methods that is applied routinely in cases of acute myeloid leukemia (AML) to assess chromosomal changes that are clinically relevant for risk classification and treatment decisions. In this study, we assess the use of Genomic Proximity Mapping® (GPM) for cytogenomic analysis of AML diagnostic specimens for detection of cytogenetic risk variants included in the European Leukemia Network (ELN) risk stratification guidelines. Archival patient samples (n = 48) from the Fred Hutchinson Cancer Center (FH) leukemia bank with historical clinical cytogenetic data were processed for GPM and analyzed with the CytoTerra cloud-based analysis platform. Genomic proximity mapping showed 100% concordance for all specific variants that have associated impacts on risk stratification as defined by ELN 2022 criteria and 78% concordance when considering all variants reported by the FH Cytogenetics Lab. Notably, the percentage of blasts (ranging from 5–96%) did not have a clear effect on the ability to detect these variants. In two cases, GPM identified a recurrent inv(9)(p13.3p13.1). These findings demonstrate GPM's effectiveness for the evaluation of known AML-associated risk variants and a source for biomarker discovery.

1 INTRODUCTION

Cytogenetic diagnostic testing is considered standard-of-care and routinely applied in acute myeloid leukemia (AML) and other hematological malignancies.[1] Decades of cytogenetic testing (including karyotyping, fluorescence *in situ* hybridization (FISH), and chromosome genomic array testing (CGAT)) have identified recurrent translocations, inversions, deletions, duplications, and copy-neutral loss of heterozygosity (cnLOH) in AML that drive tumorigenesis through oncogenic fusion genes, disrupted tumor suppressor genes, or amplified oncogenes. These biomarkers are clinically useful for assessing prognosis and guiding treatment.[2] Each cytogenetic method has specific strengths, but also limitations including low resolution (karyotyping), the need for living, dividing cells (karyotyping), limited scope of detection (FISH), and inability to detect balanced rearrangements (CGAT). These partially overlapping strengths and limitations necessitate costly multi-modal cytogenetic and molecular testing for AML that takes days to weeks to complete. Nonetheless, current risk classification schemes as well as guidelines used in AML treatment and clinical trial design require cytogenetic findings.

We have developed Genomic Proximity Mapping® (GPM), a next-generation sequencing (NGS)-based assay that uses the proximity of interacting DNA sequences in intact cells to determine the linear structure of chromosomes.[3,4] The method involves crosslinking DNA within intact nuclei (**Figure 1A**), extracting then fragmenting the crosslinked DNA, and generating chimeric DNA molecules by ligating DNA molecules that were in close three-dimensional proximity.[5] Sequences that are closer along a chromosome are more likely to be proximal and interact with each other than sequences that are distant (**Figure 1B**). Interaction frequencies identify structural variants (SVs) by detecting changes in expected interaction patterns. When visualized as a heat map, patterns of pairwise interactions can be used to interpret the structure of

SVs including translocations, inversions, insertions, deletions and duplications (**Figure 1C–1D**).[6–8] For instance, a balanced translocation is visualized as an excess of pairwise interactions between two chromosomes in a distinct pattern as illustrated in **Figure 1D**. Signal decays moving away from the breakpoint following a power law function[5] with sharp boundaries in the signal dictated by the sequences involved in the interaction. The inter-chromosomal signal is reciprocal on the heatmap in cases of balanced rearrangements, creating a 'bowtie' pattern. In cases of unbalanced rearrangements, only half of this 'bowtie' pattern is present, but the pattern on the heat map retains sharp boundaries and highest signal intensity is observed at the breakpoint. Similar patterns are observed for inversions but as intra-chromosomal signals. These patterns of long-range sequence interactions form the basis of SV detection with GPM illustrating a distinct advantage in detecting SV breakpoints with high confidence even with lower overall reads as compared to standard whole genome sequencing (WGS). The clinical applicability of GPM analysis to resolve complex chromosome abnormalities in a constitutional genetic cases has been recently published.[9,10] Importantly, GPM can be performed on freshly isolated, frozen, or formalin-fixed, paraffin-embedded patient samples making it a broadly applicable cytogenetic analysis platform.[11,12] In this study, the CytoTerra® Cytogenetics Platform was assessed for the use of GPM as a method for cytogenomic analysis of AML. Using cryopreserved diagnostic specimens, the CytoTerra platform detected all known risk variants included in the European Leukemia Network (ELN) risk stratification guidelines[2] previously identified by standard-of-care cytogenetics. Genomic proximity mapping also identified additional clinically significant variants absent from cytogenetic reports at the time of diagnosis. Furthermore, a recurrent inversion not previously observed in AML was identified in the study population of 48 samples. These observations support the effectiveness of GPM to enhance

cytogenomic characterization of AML cases.

2 METHODS

2.1. Study population and cytogenetic evaluation

The study conforms with The Code of Ethics of the World Medical Association (Declaration of Helsinki) and was approved by the FH Institutional Review Board. Written consent was obtained from all patients. Inclusion criteria included a confirmed AML diagnosis and consent for WGS. Exclusion criteria included insufficient cryopreserved samples in the FH leukemia biobank for orthogonal testing. Samples included various AML-associated genetic abnormalities identified by standard-of-care testing. Data from institutional medical records were verified manually.

2.2. CytoTerra library construction and sequencing

Thawed cryopreserved samples were counted by hemocytometer then stored in PGShield™ at 4°C. Approximately 200,000–500,000 cells were used for library preparation with the Phase Genomics CytoTerra® kit (v1.1) (Phase Genomics, Inc., Seattle, WA) following the manufacturer's protocol. In brief, 200,000 to 1 million (M) cells were crosslinked in 1% formaldehyde for 15 minutes. Following quench, cells were lysed and crosslinked chromatin was immobilized on magnetic beads. DNA was digested using restriction enzymes and the ends were filled in with biotinylated nucleotides. Bead-bound chromatin was subjected to proximity ligation and crosslinks reversed. Proximity ligated junctions were enriched using streptavidin magnetic beads. Streptavidin bead-bound fragments were used to generate an Illumina® sequencing library sequenced on a NovaSeq™ 6000 (Illumina, Inc., San Diego, CA) in the paired-end 150bp format to an average depth of 150M read pairs. Further details about the

Illumina NGS methods are provided in the **Supplemental Methods**. Library performance was evaluated using Phase Genomics' open-source quality control (QC) tool hic_qc.[13] Libraries passing primary QC metrics were advanced to analysis and include same-strand read pairs (> 20%), high-quality reads (uniquely mapping sequences, > 50%), and duplicate read pairs (< 35%).

2.3. CytoTerra Data Analysis

Raw FASTQ files were analyzed with the CytoTerra cloud-based analysis platform. First, reads were aligned to the reference genome GRCh38 using bwa mem v0.7.18[14]; duplicate reads were marked with samblaster v0.1.26; and after gathering QC statistics, duplicate reads were removed using matlock. In-house scripts filtered out low-quality alignments and calculated coverage. The Pairix v0.3.7[15] and Cooler v0.10.2[16] packages generated the contact frequency matrix. Small variants were predicted from filtered alignments using a convolutional neural network, and an in-house script extracted allele frequency data. The proprietary CytoTerra analytic suite is composed of:

- A convolutional neural network trained to detect variants in heatmap images generated from the contact frequency matrix. The neural network generates a probability confidence value for each call.
- A support vector machine caller predicts aneuploidy from coverage and allele frequency data and produces a probability confidence value for each variant.
- A bioinformatics tool specialized in breakpoint detection from GPM alignments that uses a log-odds method to produce a confidence value.
- A trio of bioinformatics tools specialized in detecting copy number and cnLOH based on alignment data and allele frequency from small variants with each reporting an Expect value to judge confidence.

After variant predictions were made, putative false positives were flagged using a rules-based approach reliant on call confidence statistics and a convolutional neural network trained in-house to classify calls based on heatmap image data. The CytoTerra Curator platform was used for variant analysis, annotation, and visualization. Calls were reviewed by a blinded operator. Variants > 100,000 bp were reported except cnLOH; while CytoTerra's resolution for cnLOH was 1 Mb, a 10 Mb threshold was employed for likely somatic cnLOH events.

3 RESULTS

3.1. Clinical AML sample performance in proximity ligation sequencing

Cryopreserved AML samples were identified from the FH AML sample repository. The study population is composed of 48 diagnostic samples derived from peripheral blood, bone marrow aspirates, and apheresis samples. Details of the patient cohort with therapy information are listed in Table 1. Blast percentage estimates varied between 5% and 96% (**Figure 2A**). Genomic proximity mapping libraries generated from 200,000–500,000 cells were successful for all samples. Post-sequencing quality control analysis of the resulting libraries showed acceptable performance for all libraries (**Figure 2B–2D**, **Supplementary Table 1**). Two normal bone marrow samples and a diverse set of normal lymphoblastic cell lines observing no SVs that met reporting criteria[17] were used as controls.

3.2. Detection of inter-chromosomal and intrachromosomal rearrangements

Based on previous cytogenetic risk assessments, the blinded CytoTerra assessment identified all ELN-identified translocations in the study population. These include the t(8;21)(q22;q22.1) translocation that generates the *RUNX1*::*RUNXT1* fusion (**Figure 3A**). Beyond the ELN-

classified translocations identified, several additional rearrangements of known and unknown significance were observed in the CytoTerra analysis. These variants included a *NUP98*::*KDM5A* fusion created by a t(11;12)(p15.4;p13.33) translocation, a variant associated with poor prognosis and chemoresistance (**Figure 3D**).[18] This rearrangement was not previously detected in the cytogenetics report for this patient. Though not previously reported in the literature, an unbalanced t(6;7)(p23;q36.3) translocation identified demonstrated a small region of non-reciprocal exchange resulting in a deletion of the *JARID2* gene, a known tumor suppressor in myeloid neoplasms (**Figure 3E**).[19] The remaining variants identified were of no known significance (**Supplementary Table 2**). We observed a single complex karyotype in this study population that included six inter-chromosomal breakpoints, in addition to a -7 and dup(21)(q22.12q22.12) (**Figure 4C**) among other less complex karyotypes (**Figure 4A–4B**).

Among this data set, inv(16)(p13.1q22) is the most common, observed in 4–5% of AML patients that receive cytogenetic work-ups.[20,21] CytoTerra identified 4/4 inv(16) rearrangements observed by cytogenetics in the study population (**Figure 5A**). The next most common class of inversions observed are related rearrangements involving the *MECOM* (*EVI1*) locus.[21] In this study, both cytogenetics and CytoTerra identified a single instance of an inv(3)(q21.3q26.1) (**Figure 5B**). However, only CytoTerra identified an inv(3)(p24.3q26.2), an unusual but previously documented rearrangement.[22] Both of these rearrangements are classified as adverse by the ELN 2022 classification.[2] We also observed an instance of an inv(12)(p13.32p13.2), a variant that has not been previously documented (**Figure 5D**).

Among the 48 cases in this set, we observed a recurrent inversion not previously documented in AML in two cases. The inv(9)(p13.3p13.1) was seen in two unrelated cases and occurs between two paralogous genes *ANKRD18A* and *ANKRD18B* (**Figure 5C**). These genes lie within regions

of segmental duplication in the pericentromeric region of chr9 at approximately 5 Mbp from each other.

3.3. Detection of insertions

In this case study populations, we identified two insertions not reported by conventional cytogenetic analysis (**Supplementary Tables 3–4**). In one case, 8 Mbp of chr5 was inserted into chr13 150 kbp upstream of *FLT3* and disrupting the *PAN3* gene (**Figure 6A)**. In another, 120 kbp of chr12 was inserted into another site on chr12 (**Figure 6B**). This small insertion is copy-neutral but does disrupt *DDX11*, which when mutated is associated with negative outcomes in AML.[23]

3.4. Detection of copy number aberrations

In the current study population, only two observations of -7 were detected by CytoTerra. Numerous aneuploidies and smaller deletions and duplications of unknown significance were detected by CytoTerra (**Supplementary Tables 4–5**). Among these are two incidences of deletions of the *TET2* gene not previously detected in cytogenetic reporting.

In the study population, a single instance of cnLOH was detected on chr13 by both CGAT and CytoTerra (**Supplementary Table 6**).

3.5. Concordance between cytogenetics and CytoTerra

Blinded review of variant calls generated by the CytoTerra platform were compared with the record of clinical cytogenetics. CytoTerra showed 100% concordance for all specified variants that have associated impacts on risk stratification as defined by ELN 2022 criteria (**Table 2**).

Notably, blast percentage did not have a clear effect on the ability to detect these variants, with % blasts ranging between 5% and 96%. When considering variants that meet the ELN categorization of "cytogenetic and/or molecular abnormalities not classified as favorable or adverse", CytoTerra demonstrated a 77.8% concordance rate with cytogenetics (**Table 2**). A majority (6/10) of discordant calls were aneuploidies (two +8, two +22, and one case of 4N, tetraploidy). Nineteen copy number and structural variant calls were discordant between GPM and clinical cytogenetics in 12 unique samples. These underwent orthogonal testing with whole genome sequencing (WGS) to corroborate their GPM or cytogenetic presentation. Fourteen GPM calls were confirmed over their cytogenetic presentation, four cytogenetic calls were corroborated over their GPM presentation, and one case with full-genome tetraploidy which neither GPM nor NGS can detect (**Supplementary Table 7**).

4 DISCUSSION

4.1. Summary

This study shows the ability of whole genome sequencing with GPM to detect cytogenetic aberrations including translocations, inversions, copy number alterations, insertions, and deletions, as well as cnLOH. Over an initial set of AML patients representing the full range of ELN prognostic risk categories, we challenged the GPM assay to recapitulate the prognostic data produced by archival clinical grade molecular and cytogenetic assays.

4.2. A role of GPM in the analysis of chromosome abnormalities

Cytogenetics encompasses at least three major technologies that are used for diagnostic and research purposes: karyotyping, FISH, and CGAT.[1] Each technology has specific strengths and

limitations, necessitating the use of all three technologies to achieve a comprehensive assessment of genomic alterations (**Table 3**). While karyotyping and CGAT both offer genome-wide assessment, karyotyping has limited resolution and CGAT is unable to identify balanced rearrangements. These limitations are likely why variants such as *NUP98::KDM5A*—a balanced rearrangement involving a sub-microscopic (~3 Mbp) sequence—eluded detection in the initial cytogenetic workup. While FISH can readily identify rearrangements such as t(11;12) that lead to the *NUP98::KDM5A* fusion,[24] it has practical limitations for the number and variety of FISH probes that can be applied to any one diagnostic sample. Genomic proximity mapping offers as an alternative that captures balanced and unbalanced alterations in a genome-wide manner, including those involving small (< 100,000 bp) amounts of sequence.

Genomic proximity mapping is one of several technologies being applied for the detection of chromosomal abnormalities for clinical research. Whole-genome sequencing (WGS) using either short-read (Illumina, Element, Ultima, etc.) or long-read (PacBio, Oxford Nanopore) is increasingly applied in clinical settings. Chromoseq™,[25] a short-read WGS assay, is currently offered through Washington University, St. Louis Department of Pathology. Though it is based on WGS, the assay only reports a defined list of recurrent SVs, likely due to the limitations of short-read sequencing to identify breakpoints produced by inversions and translocations. Sequencing-based detection of rearrangements requires that a set of reads map to the junction between the sequences participating in the rearrangement, thus demanding very high coverage data. If, as it is often the case, a rearrangement is mediated by a repetitive element,[26–28] it is usually impossible for short read sequencing to span the rearrangement junction, making these aberrations undetectable. Current estimates place the sensitivity of standard short-read sequencing for large aberration detection between 10%[29] and 70%[30] with an exceptionally high

false positive rate of up to 89%.[30–33] Genomic proximity mapping overcomes this limitation of short read sequencing by capturing ultra-long-range sequence information at relatively low sequencing depth. For instance, Chromoseq™ relies on ~1 billion read pairs of Illumina® sequencing where GPM recommends only 150M read pairs of data.

Long read sequencing technologies have made enormous progress by increasing accuracy, throughput, and reducing costs in recent years.[34,35] Long reads can help overcome the ambiguity in breakpoint identification owing to the extended sequence context, and an explosion in variant-calling algorithms supporting this activity has accompanied improvements in the technology. Despite these efforts, long-read sequencing deployment to the clinic is limited by the technical challenges of large amounts of high-molecular weight DNA needed for library construction and the cost associated with sequencing, especially in the case of somatic disease where detection of clonal variants is a requirement.[36]

Optical genome mapping (OGM) is another high-molecular weight DNA-based technology that can be used to identify SVs.[37] Rather than sequencing DNA, OGM uses sequence-specific DNA labeling to generate long DNA fragments which are electrophoresed through a capillary channel. The distance between labeled sequences on the large fragments is used to map the order and orientation of genomic sequences, analogous to restriction fragment length polymorphism mapping (RFLPs). Optical genome mapping has been applied in a systematic evaluation of AML genomes and demonstrated that, like GPM, OGM detected known variants of significance with a high degree of sensitivity.[38] Optical genome mapping struggled to identify cases of trisomy 8 (6/9 identified),[38] similar to what was observed with GPM in this study (4/6 identified, **Supplementary Table 5**). This may be in part attributed to sensitivity of these methods to detection of whole chromosome aneuploidy or bias in outgrowth of abnormal myeloid clones

during culture for karyotype analysis. Identification of cryptic rearrangements, similar to those identified in this study, has been a highlight of more recent studies supporting the utility of OGM for resolving SVs in AML patient samples.[12,39,40] A non-trivial challenge overcome in these OGM-based studies is the isolation of high-molecular weight DNA from limiting amounts of patient samples. By contrast, GPM operates on approximately 1/10 the input of OGM and is compatible with a wider variety of sample types including FFPE tissue.[10,11] GPM also has the distinct advantage of running on ubiquitous short-read sequencing platforms that have established a foothold in diagnostic laboratories and thereby increasing accessibility.

4.3. Impact of GPM on ELN risk stratification

The ELN 2022 risk classification guidelines identify a set of recurrent translocations seen in AML. Though less commonly observed than translocations, a number of recurrent inversions are known contributors to the AML phenotype. The t(8;21) translocation which results in *RUNX1*::*RUNXT1* fusion is associated with favorable outcomes and may be an important biomarker for treatment with CDK4/6 inhibitors.[41] The inv(16)(p13.1q22) inversion is the most common intra-chromosomal rearrangement and generates a CBFB::MYH11 fusion transcript, a rearrangement that portends a more favorable prognosis. The t(6;9)(p23.3;q34.1) translocation similarly generates a DEK::NUP214 fusion transcript but is associated with poor prognosis (**Figure 3B**).[42] Members of the nucleoporin gene family, including *NUP98* and *NUP214*, are known to drive AML through a variety of different partner genes that are detectible using the GPM approach. Like the nucleoporin genes, *KMT2A* is known to associate with a variety of fusion partners,[43] including the *AFDN* gene observed in this study population (**Figure 3C**). While standard-of-care cytogenetics identified both *KMT2A* and *NUP214* fusions, only

CytoTerra identified the *NUP98* fusion, a variant associated with poor outcomes.[18] In this case, CytoTerra offers benefit in identifying variants of importance over existing standard methods.

4.4. GPM discovers novel findings

Detecting insertions by traditional cytogenetics is dependent on the size and genomic content of the inserted DNA segment. Some insertions are routinely detected over the course of AML diagnosis (*e.g.*, *FLT3*-ITD), but most of the variants are anonymous and are of unknown significance. In this study, GPM detected two insertions which were not detected by cytogenetics: one 8 Mbp and one 120 Kbps in length. The resolution of GPM enables more thorough description of the insertions identified in this study, both of which are associated with genes (*FLT3, DDX11*) of known clinical importance in AML. Interestingly, a recurrent inversion was observed in two cases involving inv(9)(p13.3p13.1) between two paralogous genes *ANKRD18A* and *ANKRD18B* spanning a 5MB region and lie in a pericentromeric region and involve a sub-microscopic interval of the genome. This inversion has not been previously documented in AML genomes. Rearrangements like this can be challenging to detect by cytogenetics and standard short read sequencing methods. However, this variant has been observed previously in a number of cases of acute lymphoblastic leukemia,[44] detected only through a targeted resequencing effort of this pericentromeric region. These cytogenetically cryptic classes of rearrangement represent a relevant class of variant where CytoTerra shows promise to make a clinical impact.

Unsurprisingly, most additional variants called by CytoTerra involve copy number changes below the level of cytogenetic resolution (< 5 Mb). They may also reflect changes deemed unreportable from a clinical perspective as many of these variants are of unknown significance,

including potentially constitutional variants. However, as noted above, CytoTerra identified additional variants of known significance not specified in the ELN risk criteria, including *NUP98::KDM5A*. This, and six other translocations, and two additional inversions, were uncovered by CytoTerra in this study. Previous cytogenetic analysis likely failed to observe these rearrangements because of the genomic location and/or size of the genomic interval involved. For example, *NUP98* lies at the 11p terminus and has multiple oncogenic partners making it particularly challenging to detect.

4.5. Limitations of GPM

Relatively few copy number aberrations (CNAs) are considered informative for risk stratification by the ELN 2022 guidelines. Copy number aberrations can also be included under the catch-all category of 'cytogenetic and/or molecular abnormalities not classified as favorable or adverse', which impart intermediate risk. Other specific abnormalities are cited in the ELN guidelines including -5, del(5q), -7, -17 (or -17p), all of which are associated with adverse risk. One of GPM's limitations is its relatively lower sensitivity in detecting CNAs. This limitation may be overcome by increased depth of sequencing, allowing for higher confidence detection of subtle changes of minor allele frequency. This highlights the challenge of detecting mosaic changes in whole chromosome copy number in all sequencing-based coverage data. Tetraploidy also represents a challenge because, in the case of whole genome duplication, the allele frequency remains in balance and is undetectable by sequencing, array, or OGM. One patient demonstrated a cnLOH of 13q, a variant frequently observed in AML with *FLT3*-ITD mutations[45,46]. This singular cnLOH variant was confirmed by CGAT but more extensive studies will be necessary to determine GPM's performance detecting this class of SV. The limit of detection (LOD) has yet to

be systematically determined but anecdotal detection variants in a 5% myeloblast sample were observed in this study. In a constitutional genetics study, a variant previously estimated to be found in 7% of the patient's peripheral blood was successfully identified, consistent with a sub-10% abundance LOD.[10]

4.6. Conclusion

This study demonstrates GPM's capability to comprehensively interrogate the entire genome including detecting cryptic chromosomal aberrations at a higher resolution than conventional karyotyping and CGAT. The identification of a novel recurrent AML variant in this 48-sample study demonstrates the potential of GPM as a tool for biomarker discovery. The improved detection of ELN risk variants with GPM warrants a comprehensive study to evaluate CytoTerra for improved accuracy in patient risk stratification.

REFERENCES

1. Wan TSK. Cancer Cytogenetics: An Introduction. In: Wan TSK, Cancer Cytogenetics. Springer New York; 2017. p.1-10.

2. Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood. 2022;140(12):1345-1377.

3. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol. 2013;31(12):1119-1125.

4. Kaplan N, Dekker J. High-throughput genome scaffolding from in vivo DNA interaction frequency. Nat Biotechnol. 2013;31(12):1143-1147.

5. Lieberman-Aiden E, Van Berkum NL, Williams L, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science. 2009;326(5950):289-293.

6. Harewood L, Kishore K, Eldridge MD, et al. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. Genome Biol. 2017;18(1):125.

7. Chakraborty A, Ay F. Identification of copy number variations and translocations in cancer cells from Hi-C data. Bioinformatics. 2018;34(2):338-345.

8. Dixon JR, Xu J, Dileep V, et al. Integrative detection and analysis of structural variation in cancer genomes. Nat Genet. 2018;50(10):1388-1398.

9. Fang H, Eacker SM, Wu Y, et al. Genetic and functional characterization of inherited complex chromosomal rearrangements in a family with multisystem anomalies. Genet Med Open. 2025;3:103423.

10. Fang H, Eacker SM, Wu Y, et al. Evaluation of Genomic Proximity Mapping (GPM) for Detecting Genomic and Chromosomal Structural Variants in Constitutional Disorders. J Mol Diagn. 2025;S1525-1578(25)00194-00201.

11. Dennis MJ, Pavlick DC, Kacew A, et al. Low PD-L1 expression, MAP2K2 alterations, and enriched HPV gene signatures characterize brain metastases in head and neck squamous cell carcinoma. J Transl Med. 2024;22(1):960.

12. Gonzales PR. Integration of Newer Genomic Technologies into Clinical Cytogenetics Laboratories. Genes (Basel). 2025;16(6):688.

13. Phase Genomics. HiC_QC. *GitHub*. Available at: https://github.com/phasegenomics/hic_qc (accessed on 2024, November 1).

14. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.

15. Lee S, Vitzthum C, Alver BH, Park PJ. Pairs and Pairix: a file format and a tool for efficient storage and retrieval for Hi-C read pairs. Bioinformatics. 2022;38(6):1729-1731.

16. Abdennur N, Mirny LA. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics. 2020;36(1):311-316.

17. Ebert P, Audano PA, Zhu Q, et al. Haplotype-resolved diverse human genomes and integrated analysis of structural variation. Science. 2021;372(6537):eabf7117.

18. Mohanty S. NUP98 Rearrangements in AML: Molecular Mechanisms and Clinical Implications. Onco. 2023;3(3):147-164.

19. Celik H, Koh WK, Kramer AC, et al. JARID2 Functions as a Tumor Suppressor in Myeloid Neoplasms by Repressing Self-Renewal in Hematopoietic Progenitor Cells. Cancer Cell. 2018;34(5):741-756.e8.

20. Grimwade D, Walker H, Oliver F, et al. The Importance of Diagnostic Cytogenetics on Outcome in AML: Analysis of 1,612 Patients Entered Into the MRC AML 10 Trial. Blood. 1998;92(7):2322-2333.

21. Bendari M, Khoubila N, Cherkaoui S, et al. Current Cytogenetic Abnormalities in Acute Myeloid Leukemia. In: Aşkın Çelik T, Dey S, Chromosomal Abnormalities. London: IntechOpen; 2020. p.65-78.

22. Haferlach C, Bacher U, Grossmann V, et al. Three novel cytogenetically cryptic *EVI1* rearrangements associated with increased *EVI1* expression and poor prognosis identified in 27 acute myeloid leukemia cases. Genes Chromosomes Cancer. 2012;51(12):1079-1085.

23. Zhang T, Auer P, Dong J, et al. Whole-genome sequencing identifies novel predictors for hematopoietic cell transplant outcomes for patients with myelodysplastic syndrome: a CIBMTR study. J Hematol Oncol. 2023;16(1):37.

24. Heald JS, López AM, Pato ML, et al. Identification of novel *NUP98* fusion partners and comutations in acute myeloid leukemia: an adult cohort study. Blood Adv. 2024;8(11):2691-2694.

25. Duncavage EJ, Schroeder MC, O'Laughlin M, et al. Genome Sequencing as an Alternative to Cytogenetic Analysis in Myeloid Cancers. N Engl J Med. 2021;384(10):924-935.

26. Chen J-M, Cooper DN, Férec C, Kehrer-Sawatzki H, Patrinos GP. Genomic rearrangements in inherited disease and cancer. Semin Cancer Biol. 2010;20(4):222-233.

27. Mills RE, Walter K, Stewart C, et al. Mapping copy number variation by population-scale genome sequencing. Nature. 2011;470(7332):59-65.

28. Kidd JM, Graves T, Newman TL, et al. A Human Genome Structural Variation Sequencing Resource Reveals Insights into Mutational Mechanisms. Cell. 2010;143(5):837-847.

29. Huddleston J, Chaisson MJP, Steinberg KM, et al. Discovery and genotyping of structural variation from long-read haploid genome sequence data. Genome Res. 2017;27(5):677-685.

30. Sudmant PH, Rausch T, Gardner EJ, et al. An integrated map of structural variation in 2,504 human genomes. Nature. 2015;526(7571):75-81.

31. English AC, Salerno WJ, Reid JG. PBHoney: identifying genomic variants via long-read discordance and interrupted mapping. BMC Bioinformatics. 2014;15(1):180.

32. Tattini L, D'Aurizio R, Magi A. Detection of Genomic Structural Variants from Next-Generation Sequencing Data. Front Bioeng Biotechnol. 2015;3:92.

33. Teo SM, Pawitan Y, Ku CS, Chia KS, Salim A. Statistical challenges associated with detecting copy number variations with next-generation sequencing. Bioinformatics. 2012;28(21):2711-2718.

34. Marx V. Method of the year: long-read sequencing. Nat Methods. 2023;20(1):6-11.

35. Oehler JB, Wright H, Stark Z, Mallett AJ, Schmitz U. The application of long-read sequencing in clinical settings. Hum Genomics. 2023;17(1):73.

36. Ahsan MU, Liu Q, Perdomo JE, Fang L, Wang K. A survey of algorithms for the detection of genomic structural variants from long-read sequencing data. Nat Methods. 2023;20(8):1143-1158.

37. Mak ACY, Lai YYY, Lam ET, et al. Genome-Wide Structural Variation Detection by Genome Mapping on Nanochannel Arrays. Genetics. 2016;202(1):351-362.

38. Levy B, Baughn LB, Akkari Y, et al. Optical genome mapping in acute myeloid leukemia: a multicenter evaluation. Blood Adv. 2023;7(7):1297-1307.

39. Loghavi S, Wei Q, Ravandi F, et al. Optical genome mapping improves the accuracy of classification, risk stratification, and personalized treatment strategies for patients with acute myeloid leukemia. Am J Hematol. 2024;99(10):1959-1968.

40. Ballesta-Alcaraz L, Bernal M, Vilchez JR, et al. Application of Optical Genome Mapping for the Diagnosis and Risk Stratification of Myeloid and Lymphoid Malignancies. Int J Mol Sci. 2025;26(12):5763.

41. Swart LE, Heidenreich O. The RUNX1/RUNX1T1 network: translating insights into therapeutic options. Exp Hematol. 2020;94:1-10.

42. Slovak ML, Gundacker H, Bloomfield CD, et al. A retrospective study of 69 patients with t(6;9)(p23;q34) AML emphasizes the need for a prospective, multicenter initiative for rare 'poor prognosis' myeloid malignancies. Leukemia. 2006;20(7):1295-1297.

43. Bill M, Mrózek K, Kohlschmidt J, et al. Mutational landscape and clinical outcome of patients with de novo acute myeloid leukemia and rearrangements involving 11q23/ *KMT2A*. Proc Natl Acad Sci USA. 2020;117(42):26340-26346.

44. Sarhadi VK, Lahti L, Scheinin I, et al. Targeted resequencing of 9p in acute lymphoblastic leukemia yields concordant results with array CGH and reveals novel genomic alterations. Genomics. 2013;102(3):182-188.

45. Stirewalt DL, Pogosova-Agadjanyan EL, Tsuchiya K, Joaquin J, Meshinchi S. Copy-neutral loss of heterozygosity is prevalent and a late event in the pathogenesis of FLT3/ITD AML. Blood Cancer J. 2014;4(5):e208.

46. Gronseth CM, McElhone SE, Storer BE, et al. Prognostic significance of acquired copy-neutral loss of heterozygosity in acute myeloid leukemia. Cancer. 2015;121(17):2900-2908.

TABLES

Table 1. Clinical demographics.

Total patients	48
Sex	**n (%)**
Male	28 (58)
Female	20 (42)
Age (years)	**Mean (range)**
	54 (21-84)
GPM Specimen Source	**n (%)**
Peripheral blood	42 (88)
Bone marrow	4 (8)
Apheresis	2 (4)
Chemo/TX regimens	**n (%)**
Incomplete data	3 (6)
1	31 (65)
2-4	12 (25)
> 4	2 (4)
Transplant	**n (%)**
Yes	16 (34)
No	32 (66)

Table 2. Summary of concordance between GMP and standard-of-care cytogenetics.

Variant Class	Concordant	Discordant	Added
Specified ELN Risk Variants			
Translocation	3	0	0
Inversion	6	0	0
Copy number variant	0	0	0
Aneuploidy	3	0	0
Overall ELN	12 (100%)	0 (0%)	0
Cytogenetic and/or molecular abnormalities not classified as favorable or adverse			
Translocation	7	1	7
Inversion	7	0	2
Insertion	0	0	2
Copy number variant	5	3	29
Copy-neutral LOH	1	0	0
Aneuploidy	15	6	0
Overall	35 (77.8%)	10 (22.2%)	38

FIGURE LEGENDS

Figure 1: Principles of Genomic proximity mapping (GPM). (A) Cellular samples are collected from patients and subjected to crosslinking while still intact, freezing native chromatin conformation in place prior to proximity ligation and library generation. (B) The frequency that pairs of sequences physically interact is governed primarily by their distance along the linear length of a chromosome. Using this information, the CytoTerra variant callers can identify abnormalities in chromosome structure. (C) A visual guide to how classes of chromosome aberrations appear on the GPM sequence interaction matrix. Genomic coordinates are mirrored on X and Y axes while sequence interaction frequency is represented with increasing intensity on the heat map. Using a combination of interaction frequency and sequencing coverage depth, GPM can identify every major class of structural variation.

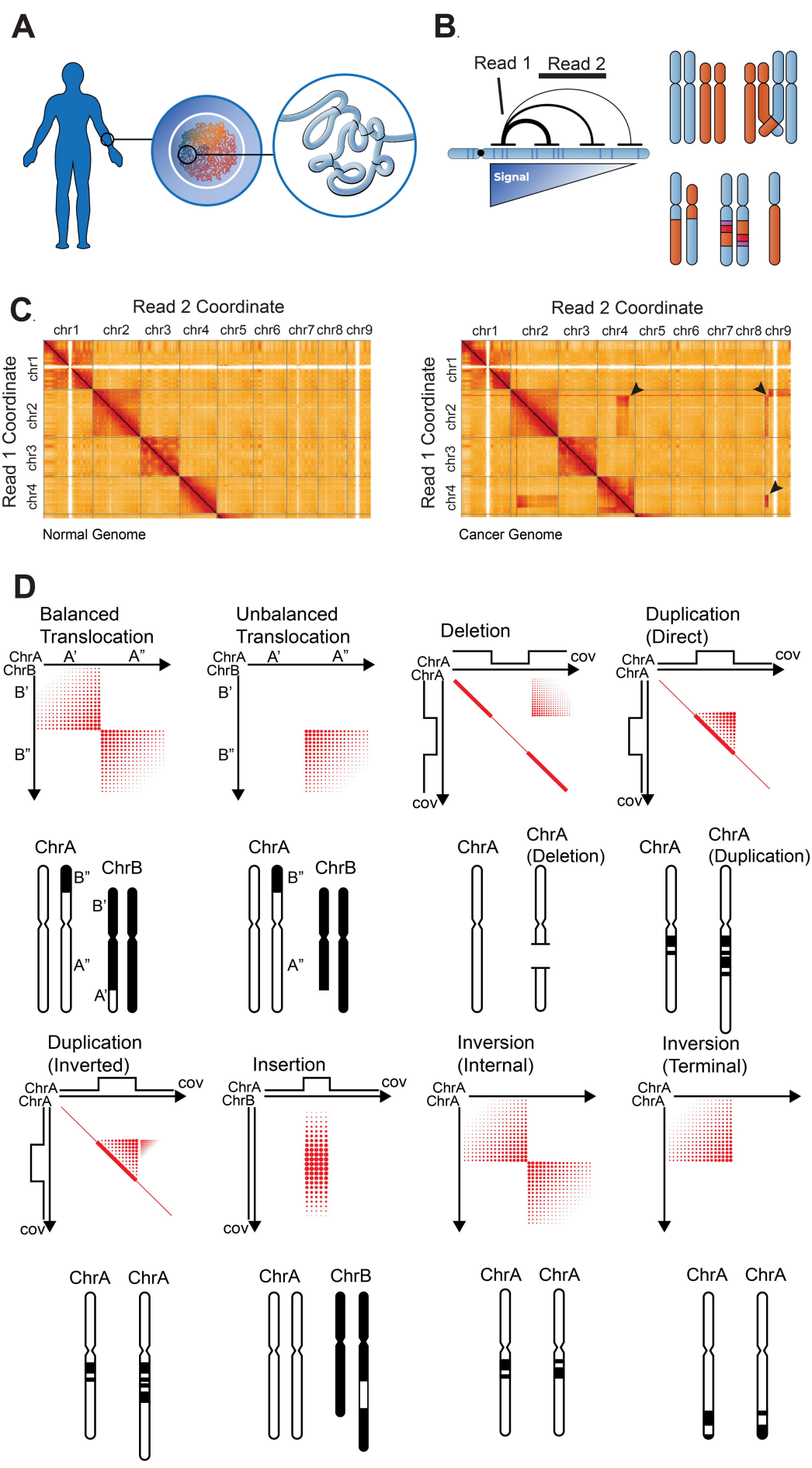
Figure 2: Blast counts and library parameters for samples used in this study. (A) Blast percentage estimates for peripheral blood (PB), bone marrow aspirates (BM), or apheresis-derived samples used for GPM library construction. (B–D) Quality control metrics for libraries generated by sample type. (B) Reads on same strand measures the percentage of read pairs from library inserts that map to the same strand of the human reference genome and are therefore the product of a proximity ligation event. Because reads and be proximity ligated to either the same or different strand configuration, the same strand percentage multiplied by 2 gives an estimate of the fraction of library fragments derived from proximity ligation events. (C) Inter-contig mapping read pairs measure the fraction of read pairs likely to be derived by spurious ligation. (D) Duplicate reads are the fraction of reads that are the result of either PCR or optical duplication.

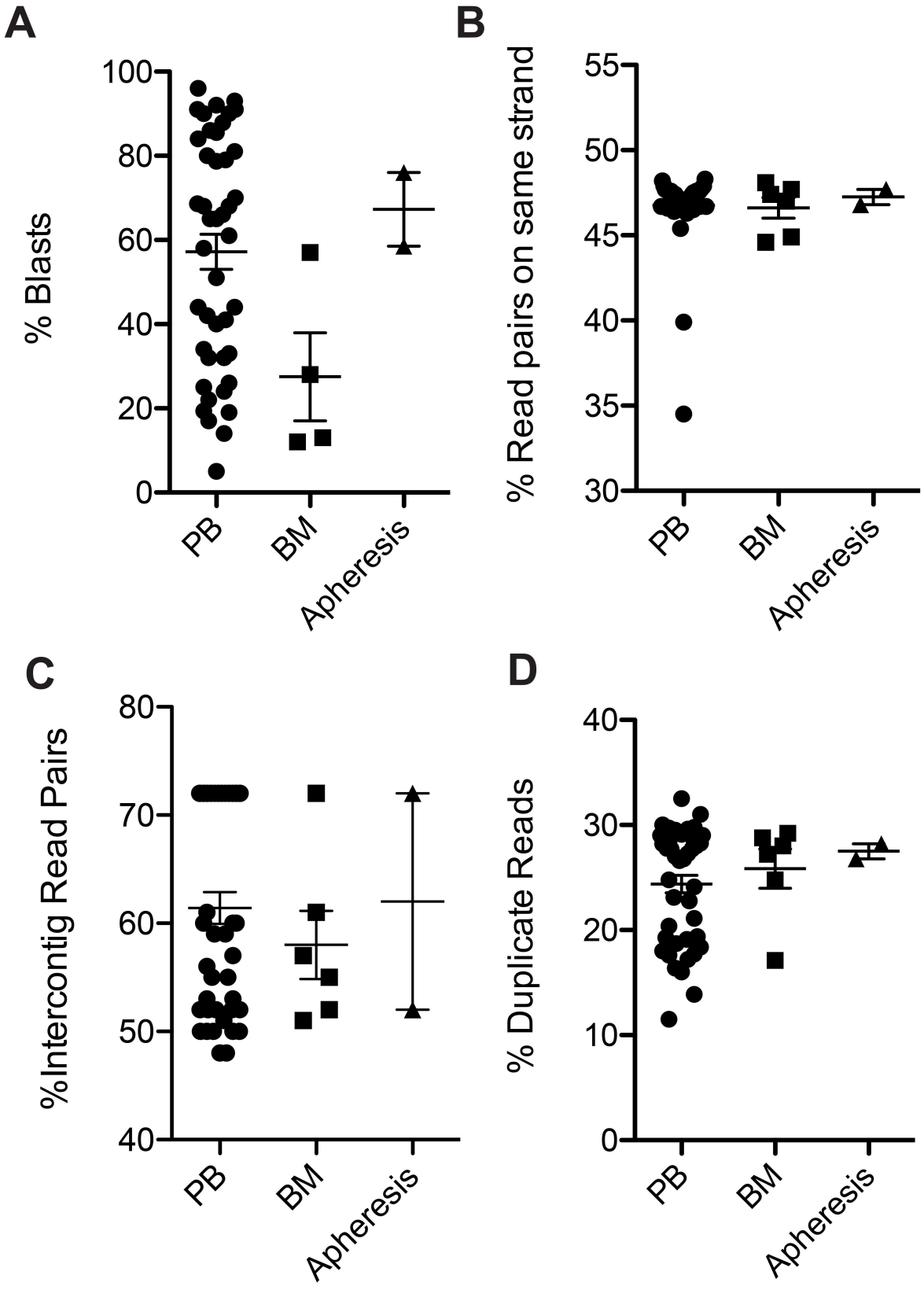
Figure 3: Selected example heatmaps of translocations detected in this study. Arrowheads indicate breakpoints observed on ideograms (black) and heatmaps (red). Coordinates of pair-wise interactions and associated gene models are labelled on X and Y axes. Gray bars indicate a lack of detected pair-wise interaction.

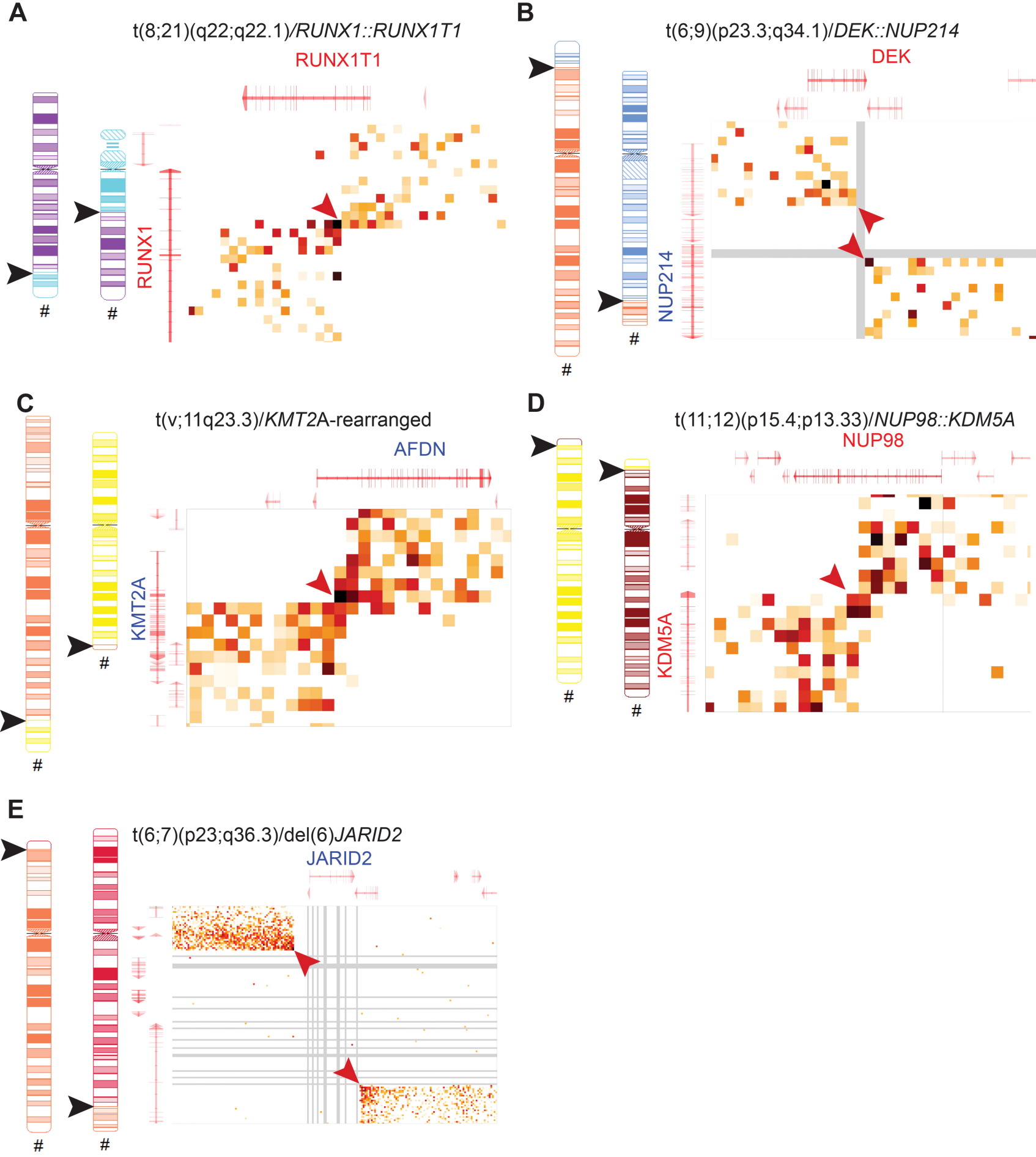
Figure 4: Circos plots illustrating the range of complexity observed in this study. (A) A normal karyotype patient sample (46, XY). (B) A patient presenting with 45,X,-Y,t(8;21)(q22;q22.1). (C) A complex series of genomic rearrangements uncovered using GPM. Inferred ISCN for this case:
45,XY,del(5)(q22q35),der(5)t(5;17)(q14.3;p13.3),-7,der(9)t(9;14)(q33.3;q23.1),der(14)del(14)(q22.2q23.1)t(9;14),der(17)t(17;18)(p13.3;q21.1),der(18)(18pter->18q21.1::14q23.1::5q35.1->5qter). Outer ring: Chromosomes represented as colored boxes, black bar illustrates the location of the centromere. Middle rings: Red line illustrates raw coverage with inferred copy number illustrated as bars below. Gray = copy 2, blue = copy 1, red > copy 2. Inner ring: minor allele frequency (MAF). Gray dots indicate expected MAF for copy 2 while red dots indicate a deviation from expected frequency.

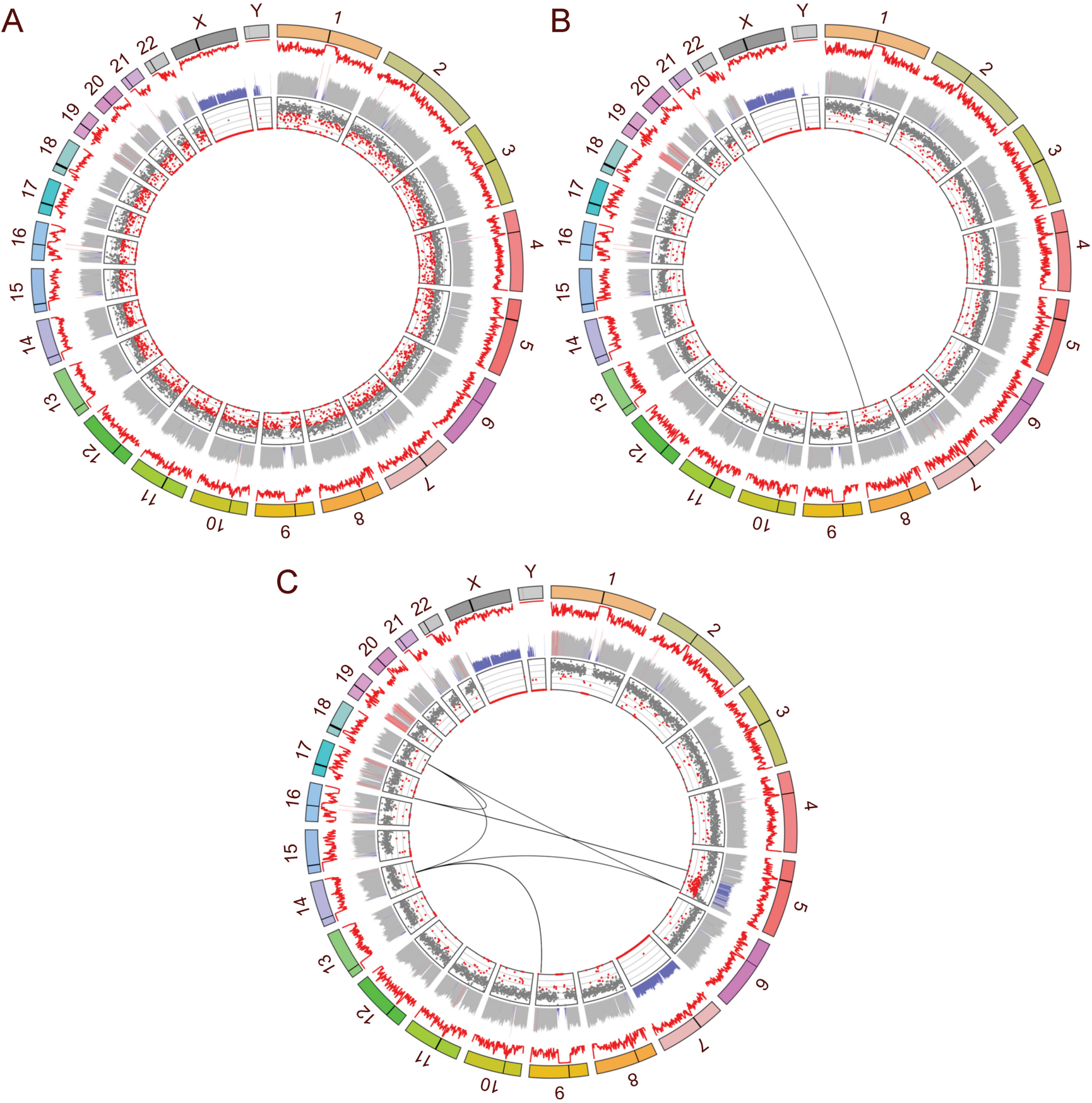
Figure 5: Selected inversions identified in the study population. (A). Example of the most common recurrent inversion observed in AML cases involving chr16, creating a *MYH11*::*CBFB* fusion gene. (B) A less common inv(3) involving *MECOM* (*EVI1*). (C) A pair of recurrent inv(9) observed in this study, a variant not previously associated with AML. Arrowheads indicate breakpoints observed on ideograms (black) and heatmaps (red).

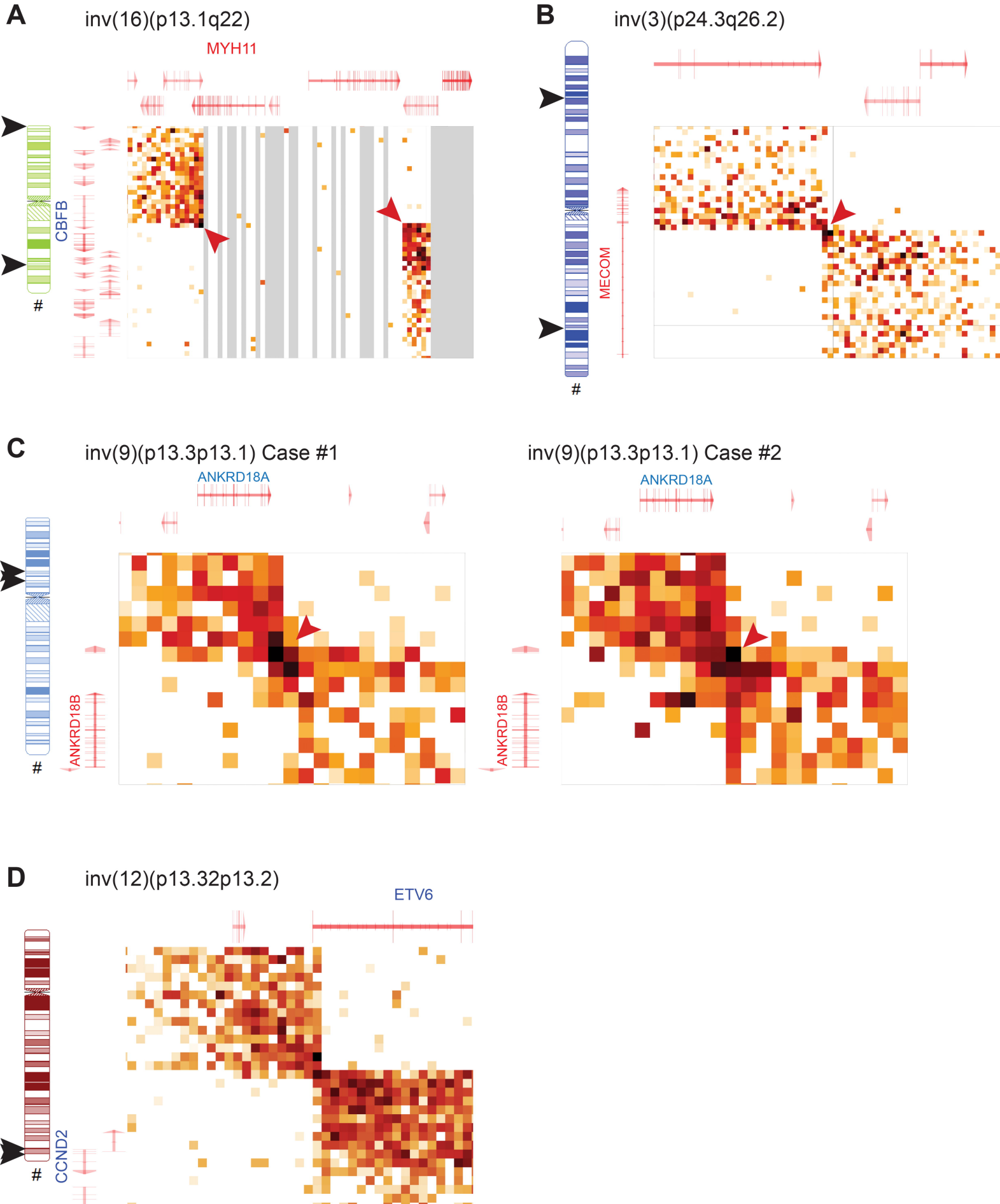
Figure 6: Two insertions observed by GPM but not reported by clinical cytogenetics. (A) Inter-chromosomal insertion: 8 Mbp of chr5 inserted into chr13 150 kbp upstream of *FLT3*, disrupting *PAN3*. (B) Intra-chromosomal insertion: 120 kbp of chr12 inserted into another region in chr12, disrupting *DDX11* (mutated *DDX11* portends poor AML prognosis). Arrowheads indicate insertion site on ideograms (black) and on heatmaps (red).

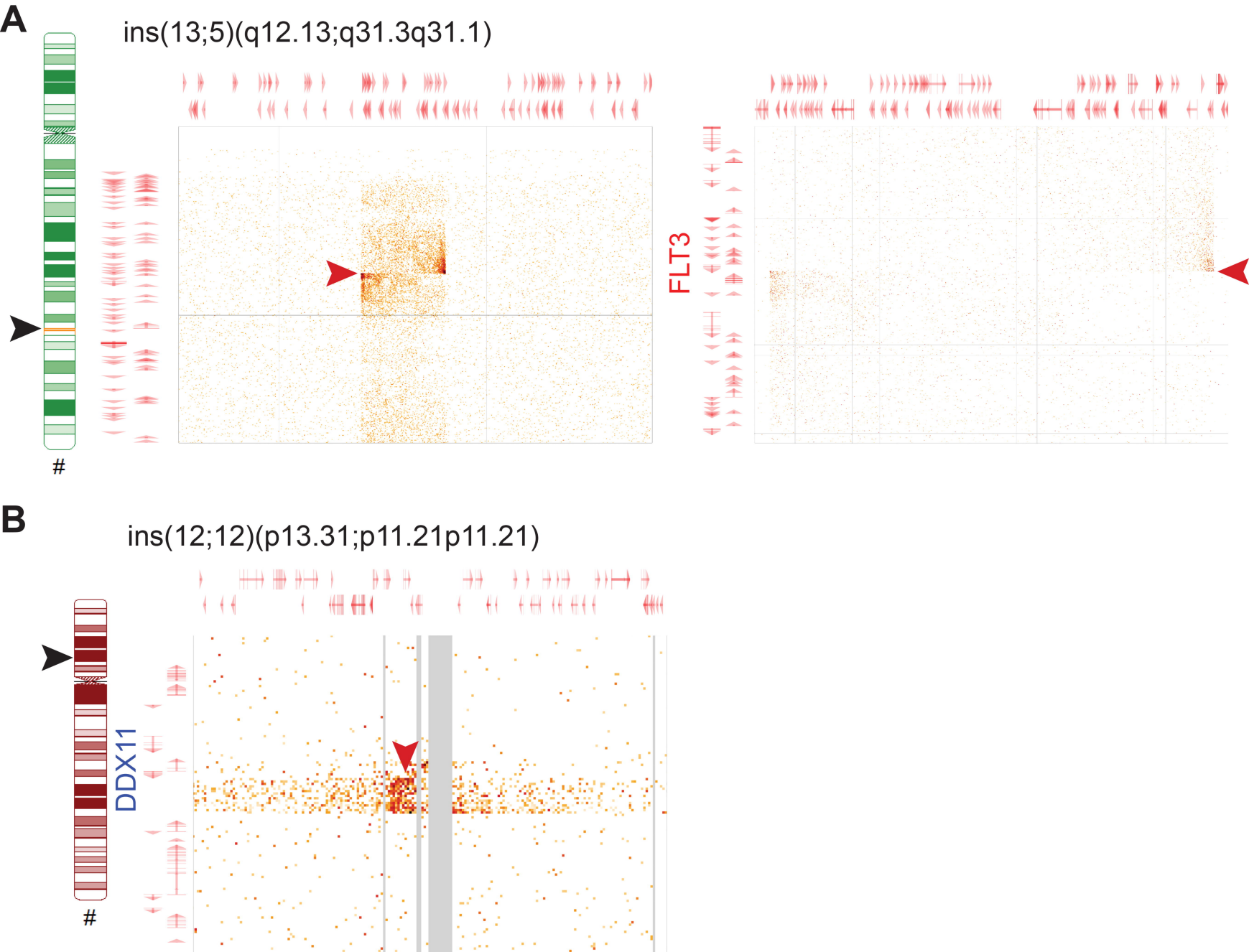

A
B
Read 1
Read 2
Signal
C
Read 2 Coordinate
chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9
Read 1 Coordinate
Normal Genome
Cancer Genome
D
Balanced Translocation
Unbalanced Translocation
Deletion
Duplication (Direct)
ChrA
ChrB
A'
A"
B'
B"
cov
Duplication (Inverted)
Insertion
Inversion (Internal)
Inversion (Terminal)
ChrA (Deletion)
ChrA (Duplication)


A
% Blasts
100
80
60
40
20
0
PB
BM
Apheresis
B
% Read pairs on same strand
55
50
45
40
35
30
PB
BM
Apheresis
C
%Intercontig Read Pairs
80
70
60
50
40
PB
BM
Apheresis
D
% Duplicate Reads
40
30
20
10
0
PB
BM
Apheresis


A
t(8;21)(q22;q22.1)/RUNX1::RUNX1T1
RUNX1T1
RUNX1
#
#
B
t(6;9)(p23.3;q34.1)/DEK::NUP214
DEK
NUP214
#
#
C
t(v;11q23.3)/KMT2A-rearranged
AFDN
KMT2A
#
#
D
t(11;12)(p15.4;p13.33)/NUP98::KDM5A
NUP98
KDM5A
#
#
E
t(6;7)(p23;q36.3)/del(6)JARID2
JARID2
#
#


A
B
C
X
Y
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22


A
inv(16)(p13.1q22)
MYH11
CBFB
#
B
inv(3)(p24.3q26.2)
MECOM
#
C
inv(9)(p13.3p13.1) Case #1
ANKRD18A
ANKRD18B
#
inv(9)(p13.3p13.1) Case #2
ANKRD18A
ANKRD18B
D
inv(12)(p13.32p13.2)
ETV6
CCND2
#


A
ins(13;5)(q12.13;q31.3q31.1)
FLT3
#
B
ins(12;12)(p13.31;p11.21p11.21)
DDX11
#

SUPPLEMENTAL METHODS

Illumina Whole Genome Sequencing

Three hundred ng of DNA from cryopreserved specimens, quantified with the Qubit 2.0 Fluorometer (ThermoFisher, Waltham, MA), were used for the study. Libraries were prepared with the Illumina DNA PCR-Free Prep, Tagmentation, and Illumina DNA/RNA UD Indexes, Tagmentation (Illumina, San Diego, CA). Library quantification was performed using the QuantStudio5 real-time PCR system from Applied Biosystems (ThermoFisher Scientific, Waltham, MA) with the KAPA Library Quantification Kit for Illumina (Roche Diagnostics Corporation, Indianapolis, IN). Equimolar concentrations of individual libraries were pooled. Sequencing was performed on the Illumina NovaSeq 6000 (Illumina, Inc, San Diego, CA) with a paired-end, 150-base read length sequencing configuration. Four S4-300 flow cells were employed. The average sequencing output per library was 861.2.4M read pairs (range 595M-1124.8M).

Data analysis was conducted with the DRAGEN Somatic analysis pipeline v4.3.6 in tumor-only mode, implemented in BaseSpace (Illumina, San Deigo, CA). The hg38 human genome assembly was used as a reference. Illumina's filters were used to filter artifacts in the analysis: systematic noise filters for single nucleotide variants (SNV SomaticSystematic Noise v2.0.0) and structural variants (SV Systematic Noise Baseline Collection v3.0.0). Nirvana Variant Annotation and DUX4 rearrangement callers were implemented.

Supplementary Table 1. Post-sequencing quality control results. Post-sequencing quality control analysis of GPM library preparations showed acceptable performance for all libraries.

Sample #	% duplicate reads	% intercontig pairs HQ	% pairs on same strand HQ	Total read pairs	Total read pairs HQ
1	0.268	0.3843	0.4683	128138316	66756524
2	0.1595	0.3485	0.4671	158198178	97018854
3	0.2919	0.2957	0.4807	255791964	130213089
4	0.3254	0.3129	0.4753	171906050	82798933
5	0.2977	0.383	0.4767	139929729	70352432
6	0.2774	0.3034	0.4819	308417004	160210996
7	0.2896	0.2662	0.476	187717273	95437230
8	0.2996	0.31	0.4717	306884855	155358240
9	0.2114	0.2396	0.466	284307270	161045018
10	0.2278	0.3071	0.4669	261342993	148199028
11	0.1929	0.4692	0.4536	238220727	134518042
12	0.2407	0.303	0.4714	408779334	228990767
13	0.1391	0.6086	0.3449	311895105	194610052
14	0.1912	0.2924	0.4671	194916399	115308883
15	0.2035	0.2787	0.4699	233983297	135128096
16	0.1866	0.2816	0.4682	194097094	116218179
17	0.2306	0.3166	0.464	234451225	131271273
18	0.1145	0.516	0.3991	74235391	47263475
19	0.1719	0.4902	0.4716	143506049	86668251
20	0.1774	0.3201	0.4765	180468720	109404100
21	0.284	0.3621	0.4737	141391492	74694588
22	0.2724	0.4978	0.4701	288961236	150095993
23	0.2776	0.3227	0.475	195732656	100742735
24	0.28	0.3703	0.469	89535264	45514809
25	0.2959	0.3684	0.4756	213516744	107490569
26	0.2874	0.3781	0.4759	61608272	31941310
27	0.2483	0.2857	0.4793	117674331	62947370
28	0.2826	0.3609	0.4695	147586536	77203129
29	0.277	0.3551	0.4717	185706559	97588610
30	0.2899	0.3279	0.4791	192816245	100354469
31	0.2968	0.3682	0.477	195825713	98791628
32	0.3101	0.3761	0.4713	359911424	182650007
33	0.1638	0.3246	0.4756	111336166	69489188
34	0.1763	0.2911	0.4709	255236310	153930933
35	0.2823	0.386	0.4775	308994411	157363877
36	0.2843	0.33	0.4826	197568214	100152984
37	0.2681	0.533	0.4627	202344451	106492243
38	0.2728	0.3567	0.4759	284225164	150362784
39	0.2802	0.4159	0.4736	107191740	55762051
40	0.2696	0.3847	0.4745	292759260	152203548
41	0.286	0.4191	0.4698	356567893	181951672
42	0.2947	0.3826	0.4724	241731837	122552004

43	0.282	0.3379	0.4767	232292327	118715458
44	0.2905	0.3671	0.4764	296885773	151746852
45	0.2884	0.3312	0.4769	242870427	123639428
46	0.1942	0.3506	0.4645	231639007	134835553
47	0.18	0.3232	0.4709	174546387	103768875
48	0.184	0.205	0.4737	188608507	111095087

Abbreviations: GPM: Genomic Proximity Mapping™

Supplemental Table 2. Inter-chromosomal variants identified by GPM compared with the corresponding clinical cytogenetic presentation. "Discordant" means the given abnormality was reported by clinical cytogenetics but missed or reported differently by GPM. "Added" means the given variant was reported by GPM but not by clinical cytogenetics.

Patient #	GPM findings	Clinical cytogenetics ISCN	Structural Aberrations	ELN class-defining variant?	Concordance between GPM and clinical cytogenetics
13	gpm 45,X,t(8;21)(q21.3;q22.12),dup(10)(q21.1)[0.6]	45,X,-Y,t(8;21)(q22;q22)[19]/46,XY[1].nuc ish(RUNX1T1,RUNX1)x3(RUNX1T1 con RUNX1x2)[180/200]	t(8;21)(q21.3;q22.12)	Yes	Concordant
18	gpm 46,XX,del(6)(p24.1p24.1),t(6;9)(p22.3;q34.12),t(6;15)(p24.2p24.1;q21.1),dup(9)(p21.1p13.3)[0.3]	46,XX,t(6;9)(p23;q34)[3]/46,sl,t(6;15)(p23;q21)[5]/47,sdl,+13[10]/46,XX[2]	t(6;9)(p22.3;q34.12)	Yes	Concordant
			t(6;15)(p24.2p24.1;q21.1)	No	Concordant
32	gpm 46,XX,t(6;11)(q27;q23.3),dup(9)(q32)	46,XX,t(6;11)(q27;q23)[20].nuc ish(MLLx2)(5'MLL sep 3'MLLx1)[152/200]	t(6;11)(q27;q23.3)	Yes	Concordant
8	gpm 46,XX,del(6)(p23p22.3),t(6;7)(p23;q36.3),del(7)(q36.3),inv(12)(p13.32p13.2)	46,XX,t(6;7)(p2?2;q36)[20].arr[GRCh38] 6p23p22.3(15,138,707_15,564,800)x1[0.85],7q36.3(157,093,285_157,944,041)x1[0.85]	t(6;7)(p23;q36.3	No	Concordant
26	gpm 46,XX,t(10;17)(p11.2;q11.2)*[0.2]	46,XX,add(17)(p13)[2]/46,sl,del(10)(q24)[2]/46,XX,t(10;17)(p10;p10)[5]/46,XX[11].nuc ish(TP53,CEP17)x2[200]	t(10;17)(p10;p10)	No	Discordant
24	gpm 46,XY	46,XY,t(2;19)(q35;p13.3)[5]/45,XY,-21[3]/46,XY[12]	t(2;19)(q35;p13.3)	No	Discordant
40	gpm 51,XY,+4,+8,t(11;12)(p15.4;p13.33),+12,+16[0.6] [0.8]	51,XY,+4,+6,+8,+12,+16[13]/46,XY[7]	t(11;12)(p15.4;p13.33)	No	Added
42	gpm	47,XY,+X[4]/46,XY[16]	ins(13;5)(q12.13;q31.3q31.1)	No	Added

	46,XY,dup(3)(q26.31q26.31),ins(13;5)(q12.13;q31.3q31.1)[0.5]				
48	gpm 45,XY,t(5;14)(q35.1;q23.1),t(5;17)(q14.3;p13.3),t(5;18)(q35.1;q21.1),-7,t(9;14)(q33.3;q23.1),t(14;18)(q23.1;q21.1),t(17;18)(p13.3;q21.1),dup(21)(q22.12)[0.5]	45,XY,-7[14]/45,sl,del(5)(q22q35)[4]/44,sdl,t(X;9)(p11.2;p22),t(2;11)(p21;q13),?inv(10)(p11.2q11.2),del(13)(q14q31),der(13;22)(q10;q10)[2],nuc ish(D5S23x2,EGR1x1)[128/200],(D7Z1,D7S486)x1[190/200]	t(5;17)(q14.3;p13.3)	No	Added
			t(5;18)(q35.1;q21.1)	No	Added
			t(5;14)(q35.1;q23.1)	No	Added
			t(9;14)(q33.3;q23.1)	No	Added
			t(14;18)(q23.1;q21.1)	No	Added
			t(17;18)(p13.3;q21.1)	No	Added

* Retrospectively observed

Note: Blank rows indicate longitudinal cases within the histories of the patient represented in the most recent occupied row.

Abbreviations: GPM: Genomic Proximity Mapping™, ISCN: International System for Human Cytogenomic Nomenclature, ELN: European Leukemia Network

Supplementary Table 3. Intra-chromosomal abnormalities detected by GPM compared with the corresponding clinical cytogenetic presentation. "Discordant" means the given abnormality was reported by clinical cytogenetics but missed or reported differently by GPM. "Added" means the given variant was reported by GPM but not by clinical cytogenetics.

Patient #	GPM findings	Clinical cytogenetics ISCN	Inversion	ELN class-defining variant?	Other Risk	Concordance between GPM and clinical cytogenetics
2	gpm 46,XY,inv(16)(p13.11q22.1)[0.8]	46,XY,inv(16)(p13.1q22)[13]/92<4N>,slx2[7].nuc ish(CBFBx2)(5'CBFB sep 3'CBFBx1)[119/200]/(CBFBx4)(5'CBFB sep 3'CBFBx2)[44/200]	inv(16)(p13.11q22.1)	Yes		Concordant
7	gpm 45,XY,inv(3)(p24.3q26.2),-7[0.8]	45,XY,-7[11]/46,XY[9]	inv(3)(p24.3q26.2)	Yes		Added
16	gpm 46,XY,del(7)(q22.1q36.1),dup(8)(p23.3),dup(12)(p13.31),inv(16)(p13.11q22.1) [0.8]	46,XY,del(7)(q22q36),inv(16)(p13.1q22)[17]/47,XY,inv(16)(p13.1q22),+22[3].nuc ish(D7Z1x2,D7S486x1)[189/200]	inv(16)(p13.11q22.1)	Yes		Concordant
38	gpm 46,XX,inv(16)(p13.11q22.1)[0.8]	46,XX,inv(16)(p13.1q22)[16]/46,XX[4]	inv(16)(p13.11q22.1)	Yes		Concordant
47	gpm 45,XY,der(3)del(3)(q21.3)inv(3)(q21.3q26.1)del(3)(q26.1q26.2),dup(4)(q32.2q32.2)?c,-7	45,XY,inv(3)(q21q26.2),-7[16]/45,sl,del(6)(p23)[4]	inv(3)(q21.3q26.1)	Yes		Concordant
8	gpm 46,XX,del(6)(p23p22.3),t(6;7)(p23;q36.3),del(7)(q36.3),inv(12)(p13.32p13.2)	46,XX,t(6;7)(p2?2;q36)[20].arr[GRCh38] 6p23p22.3(15,138,707_15,564,800)x1[0.85],7q36.3(157,093,285_157,944,041)x1[0.85]	inv(12)(p13.32p13.2)	No	No	Added
9	gpm 46,XY,dup(1)(p36.22p36.22)?c,inv(9)(p13.3p13.1)	46,XY[20].arr[GRCh38]13q12.13qter(27,055,669_114,338,054)x2 hmz[0.75]	inv(9)(p13.3p13.1)	No	No	Added
15	gpm 48,XX,+8,inv(9)(p13.3p13.1),ins(12;12)(p13.31;p11.21p11.21),inv(16)(p13.11q22.1),+22	48,XX,+8,inv(16)(p13.1q22),+22[20]	inv(9)(p13.3p13.1)	No	No	Added
			ins(12)(9320000-9390000)<-(12)(31140000-31260000)	No	Yes	Added

	inv(16)(p13.11q22.1)	Yes	Concordant

* Retrospectively observed

Note: Blank rows indicate longitudinal cases within the histories of the patient represented in the most recent occupied row.

Abbreviations: GPM: Genomic Proximity Mapping™, ISCN: International System for Human Cytogenomic Nomenclature, ELN: European Leukemia Network

Supplementary Table 4. Copy number variant (CNV) abnormalities detected by GPM compared with the corresponding clinical cytogenetic presentation. "Discordant" means the given abnormality was reported by clinical cytogenetics but missed or reported differently by GPM. "Added" means the given variant was reported by GPM but not by clinical cytogenetics.

Patient #	GPM findings	Clinical cytogenetics ISCN	CNV	ELN class-defining variant?	Other Risk	Concordance between GPM and clinical cytogenetics
6	gpm 46,XX,dup(7)(q35q35)?c	46,XX[20]	dup(7)(q35q35)	No	No	Added
8	gpm 46,XX,del(6)(p23p22.3),t(6;7)(p23;q36.3),del(7)(q36.3),inv(12)(p13.32p13.2)	46,XX,t(6;7)(p2?2;q36)[20].arr[GRCh38] 6p23p22.3(15,138,707_15,564,800)x1[0.85],7q36.3(157,093,285_157,944,041)x1[0.85]	del(6)(p23p22.3)	No	No	Concordant
			del(7)(q36.3)	No	No	Concordant
9	gpm 46,XY,dup(1)(p36.22p36.22)?c,inv(9)(p13.3p13.1)	46,XY[20]	dup(1)(p36.22p36.22)	No	No	Added
11	gpm 46,XY,del(4)(q24q24)	47,XY,+22[2]/46,XY[18]	del(4)(q24q24)	No	Yes	Added
12	gpm 46,XY,del(4)(q24q24)[0.8]	46,XX[20]	del(4)(q24q24)	No	Yes	Added
13	gpm 45,X,t(8;21)(q21.3;q22.12),dup(10)(q21.1q21.1)[0.6]	45,X,-Y,t(8;21)(q22;q22)[19]/46,XY[1].nuc ish(RUNX1T1,RUNX1)x3(RUNX1T1 con RUNX1x2)[180/200]	dup(10)(q21.1)	No	No	Added
14	gpm 46,XY,dup(2)(p22.3p22.3)?c	46,XY[20].arr[GRCh38] 6pterp22.1(1_30,006,723)x2hmz[0.9]	dup(2)(p22.3p22.3)	No	No	Added
16	gpm 46,XY,del(7)(q22.1q36.1),dup(8)(p23.3p23.3),dup(12)(p13.31p13.31),inv(16)(p13.11q22.1) [0.8]	46,XY,del(7)(q22q36),inv(16)(p13.1q22)[17]/47,XY,inv(16)(p13.1q22),+22[3].nuc ish(D7Z1x2,D7S486x1)[189/200]	del(7)(q22.1q36.1)	No	No	Concordant
			dup(8)(p23.3p23.3)	No	No	Added
			dup(12)(p13.31p13.31)	No	No	Added

17	gpm 46,XY,dup(17)(q25.1q25.1)[0.6]	46,XY[20].arr[GRCh38] 3p21.31(46,531,828_47,287,254)x1[0.65],9q21.32(83,766,899_84,002,350)x1[0.65]	dup(17)(q25.1q25.1)	No	No	Added
			3p21.31(46,531,828_47,287,254)	No	No	Discordant
			9q21.32(83,766,899_84,002,350)	No	No	Discordant
18	gpm 46,XX,del(6)(p24.1p24.1),t(6;9)(p22.3;q34.12),t(6;15)(p24.2p24.1;q21.1),dup(9)(p21.1p13.3)[0.3]	46,XX,t(6;9)(p23;q34)[3]/46,sl,t(6;15)(p23;q21)[5]/47,sdl,+13[10]/46,XX[2]	del(6)(p24.1)	No	No	Added
			dup(9)(p21.1p13.3)	No	No	Added
20	gpm 46,XY,del(9)(q21.11q31.1)	46,XY,del(9)(q22;q34)[33]/46,idem,t(6;14)(q21;q32)[2] ** on a different day				FALSE
25	gpm 47,XY,+13,dup(15)(q13.3q13.3)[0.8]	47,XY,+13[16]/46,XY[4]	dup(15)(q13.3q13.3)	No	No	Added
26	gpm 46,XX,t(10;17)(p11.2;q11.2)*[0.2]	46,XX,add(17)(p13)[2]/46,sl,del(10)(q24)[2]/46,XX,t(10;17)(p10;p10)[5]/46,XX[11]. nuc ish(TP53,CEP17)x2[200]	del(10)(q24)[2]	No	No	Discordant
27	gpm 46,XX,del(2)(p23.3p23.3),del(16)(p13.11p12.3),(13)x2 hmz	46,XX[20].arr[GRCh37] 2p23.3(24,587,652_26,417,829)x1, 13q12.11qter(19,814,912_115,103,529)x2 hmz, 16p13.13p12.3(12,040,511_18,539,704)x1	del(2)(p23.3p23.3)	No	No	Concordant
			del(16)(p13.13p12.3)	No	No	Concordant
29	gpm 46,XY,dup(11)(q22.3q22.3)?c	46,XY[20].nuc ish(MECOMx2)[200],(DEK,NUP214)x2[200],(MLLx2)[200].arr(1-22)x2,(X,Y)x1	dup(11)(q22.3q22.3)?c	No	No	Added

30	gpm 46,XY,del(2)(p23.3p23.3)[0.8]	46,XY[20].nuc ish(MECOMx2)[200],(DEK,NUP214)x2[200],(RUNX1T1,RUNX1)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200],(RARAx2)[200].arr[GRCh37] 2p23.3(24,190,632_25,989,981)x1	del(2)(p23.3)	No	No	Concordant
31	gpm 46,XX,dup(12)(p13.31p13.31)?c	46,XX[20]	dup(12)(p13.31p13.31)?c	No	No	Added
32	gpm 46,XX,t(6;11)(q27;q23.3),dup(9)(q32q32)	46,XX,t(6;11)(q27;q23)[20].nuc ish(MLLx2)(5'MLL sep 3'MLLx1)[152/200]	dup(9)(q32)	No	No	Added
34	gpm 46,XX	46,XX,del(7)(q31)[5]/46,XX[15]	del(7)(q31)[5/20]	No	No	Discordant
37	gpm 46,XY,dup(3)(q26.31q26.31)[0.3]	46,XY[20].nuc ish(MECOMx2)[200],(DEK,NUP214)x2[200],(RUNX1T1,RUNX1)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200].arr(1-22)x2,(X,Y)x1	dup(3)(q26.31)	No	No	Added
41	gpm 46,XY,dup(20)(p13p13)?c	46,XY[20].nuc ish(MECOMx2)[200],(DEK,NUP214)x2[200],(RUNX1T1,RUNX1)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200].arr(1-22)x2,(X,Y)x1	dup(20)(p13p13)	No	No	Added
42	gpm 46,XY,dup(3)(q26.31q26.31),ins(13;5)(q12.13;q31.3q31.1)[0.5]	47,XY,+X[4]/46,XY[16]	dup(3)(q26.31q26.31)	No	No	Added
44	gpm 46,XX,dup(11)(q14.2q14.3)?c	46,XX[20].nuc ish(MECOMx2)[200],(DEK,NUP214)x2[200],(RUNX1T1,RUNX1)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200].arr(1-	dup(11)(q14.2q14.3)?c	No	No	Added

		22,X)x2				
45	gpm 46,XX,dup(3)(q26.31q26.31)?c,dup(8)(q12.1q12.1)?c	46,XX[20].nucish(MECOMx2)[200],(DEK,NUP214)x2[200],(RUNX1T1,RUNX1)x2[200],(ABL1,ASS1,BCR)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200]	dup(3)(q26.31q26.31)	No	No	Added
			dup(8)(q12.1q12.1)	No	No	Added
46	gpm 47,XY,dup(1)(q41q41)?c,dup(12)(p13.33p13.33)?c,del(19)(q13.11q13.12),+21	47,XY,+21c[20].nucish(MECOMx2)[200],(D5S23,EGR1)x2[200],(DEK,NUP214)x2[200],(D7Z1,D7S486)x2[200],(RUNX1T1x2,RUNX1x3)[147/200],(ABL1,ASS1,BCR)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200],(TP53,CEP17)x2[200].arr[GRCh37]19q13.11q13.12(33,503,646_37,428,465)x1[0.95],(21)x3c	dup(1)(q41q41)?c	No	No	Added
			dup(12)(p13.33p13.33)?c	No	No	Added
			del(19)(q13.11q13.12)	No	No	Concordant
47	gpm 45,XY,der(3)del(3)(q21.3)inv(3)(q21.3q26.1)del(3)(q26.1q26.2),dup(4)(q32.2q32.2)?c,-7	45,XY,inv(3)(q21q26.2),-7[16]/45,sl,del(6)(p23)[4]	del(3)(q21.3)	No	No	Added
			del(3)(q26.1q26.2)	No	No	Added
			dup(4)(q32.2)	No	No	Added
			del(6)(p23)[4/20]	No	No	Discordant
48	gpm 45,XY,t(5;14)(q35.1;q23.1),t(5;17)(q14.3;p13.3),t(5;18)(q35.1;q21.1),-7,t(9;14)(q33.3;q23.1),t(14;18)(q23.1	45,XY,-7[14]/45,sl,del(5)(q22q35)[4]/44,sdl,t(X;9)(p11.2;p22),t(2;11)(p21;q13),?inv(10)(p11	dup(21)(q22.12)	No	No	Added

;q21.1),t(17;18)(p13.3;q21.1),dup(21)(q22.12)[0.5]	.2q11.2),del(13)(q14q31),der(13;22)(q10;q10)[2].nuc ish(D5S23x2,EGR1x1)[128/200],(D7Z1,D7S486)x1[190/200]			
		del(5)(q22q35)	Yes	Concordant

* Retrospectively observed

** Due to complex multi-chromosomal unbalanced translocations

Note: Blank rows indicate longitudinal cases within the histories of the patient represented in the most recent occupied row.

Abbreviations: GPM: Genomic Proximity Mapping™, ISCN: International System for Human Cytogenomic Nomenclature, ELN: European Leukemia Network, CNV: copy number variation

Supplementary Table 5. Aneuploidy variants detected by GPM compared with the corresponding clinical cytogenetic presentation. "Discordant" means the given abnormality was reported by clinical cytogenetics but missed or reported differently by GPM. "Added" means the given variant was reported by GPM but not by clinical cytogenetics.

Patient #	GPM findings	Clinical cytogenetics ISCN	Aneuploidy	ELN class-defining variant?	Other Risk	Concordance between GPM and clinical cytogenetics
7	gpm 45,XY,inv(3)(p24.3q26.2),-7[0.8]	45,XY,-7[11]/46,XY[9]	-7[11/20]	Yes		Concordant
47	gpm 45,XY,der(3)del(3)(q21.3)inv(3)(q21.3q26.1)del(3)(q26.1q26.2),dup(4)(q32.2q32.2)?c,-7	45,XY,inv(3)(q21q26.2),-7[16]/45,sl,del(6)(p23)[4]	-7[20/20]	Yes		Concordant
48	gpm 45,XY,t(5;14)(q35.1;q23.1),t(5;17)(q14.3;p13.3),t(5;18)(q35.1;q21.1),-7,t(9;14)(q33.3;q23.1),t(14;18)(q23.1;q21.1),t(17;18)(p13.3;q21.1),dup(21)(q22.12)[0.5]	45,XY,-7[14]/45,sl,del(5)(q22q35)[4]/44,sdl,t(X;9)(p11.2;p22),t(2;11)(p21;q13),?inv(10)(p11.2q11.2), del(13)(q14q31),der(13;22)(q10;q10)[2]	-7[20/20]	Yes		Concordant
2	gpm 46,XY,inv(16)(p13.11q22.1)	46,XY,inv(16)(p13.1q22)[13]/92<4N>,slx2[7].nuc ish(CBFBx2)(5'CBFB sep 3'CBFBx1)[119/200]/(CBFBx4)(5'CBFB sep 3'CBFBx2)[44/200]	tetraploid[7/20]	No		Discordant
3	gpm 46,XY	47,XY,+8[15]/46,XY[5]	+8[15/20]	No	No	Discordant
5	gpm 46,XY	47,XY,+8[5]/45,X,-Y[3]/46,XY[12]	+8[5/20]/-Y[3/20]	No	No	Discordant
11	gpm 46,XY,del(4)(q24q24)	47,XY,+22[2]/46,XY[18]	+22[2/20]	No	No	Discordant
15	gpm 48,XX,+8,inv(9)(p13.3p13.1),ins(12;12)(p13.31;p11.21p11.21),inv(16)(p13.11q22.1),+22	48,XX,+8,inv(16)(p13.1q22),+22[20]	+8,+22[20/20]	No	No	Concordant
16	gpm 46,XY,del(7)(q22.1q36.1),dup(8)(p23.3),dup(12)(p13.31),inv(16)(p13.11q22.1) [0.8]	46,XY,del(7)(q22q36),inv(16)(p13.1q22)[17]/47,XY,inv(16)(p13.1q22),+22[3]	+22[3/20]	No	No	Discordant
18	gpm	46,XX,t(6;9)(p23;q34)[3]/46,sl,t(6;15)(p23;q2	+13[10/20]	No	No	Discordant

	46,XX,del(6)(p24.1p24.1),t(6;9)(p22.3;q34.12),t(6;15)(p24.2p24.1;q21.1),dup(9)(p21.1p13.3)[0.3]	1)[5]/47,sdl,+13[10]/46,XX[2]				
19	gpm 47,XY,+8	47,XY,+8[4]/46,sl,-Y[11]/46,XY[5]	+8[15/20]	No	No	Concordant
			-Y[11/20]	No	No	Discordant
24	gpm 46,XY	46,XY,t(2;19)(q35;p13.3)[5]/45,XY,-21[3]/46,XY[12]	-21[3/20]	No	No	Discordant
25	gpm 47,XY,+13,dup(15)(q13.3q13.3)[0.8]	47,XY,+13[16]/46,XY[4]	+13[16/20]	No	No	Concordant
35	gpm 47,XY,+8,(13)x2 hmz[0.8]	47,XY,+8[16]/46,XY[4].nuc ish(D8Z2x3)[161/200]	+8[16/20]	No	No	Concordant
40	gpm 51,XY,+4,+8,t(11;12)(p15.4;p13.33),+12,+16[0.6] [0.8]	51,XY,+4,+6,+8,+12,+16[13]/46,XY[7]	+4[13/20]	No	No	Concordant
			+6[13/20]	No	No	Concordant
			+8[13/20]	No	No	Concordant
			+12[13/20]	No	No	Concordant
			+16[13/20]	No	No	Concordant
46	gpm 47,XY,dup(1)(q41q41)?c,dup(12)(p13.33p13.33)?c,del(19)(q13.11q13.12),+21	47,XY,+21c[20].nuc ish(MECOMx2)[200],(D5S23,EGR1)x2[200],(DEK,NUP214)x2[200],(D7Z1,D7S486)x2[200],(RUNX1T1x2,RUNX1x3)[147/200],(ABL1,ASS1,BCR)x2[200],(MLLx2)[200],(PML,RARA)x2[200],(CBFBx2)[200],(TP53,CEP17)x2[200].arr[GRCh37] 19q13.11q13.12(33,503,646_37,428,465)x1[0.95],(21)x3c	+21[20/20]	No	No	Concordant
18	gpm 46,XX,del(6)(p24.1p24.1),t(6;9)(p22.3;q34.12),t(6;15)(p24.2p24.1;q21.1),dup(9)(p21.1p13.3)[0.3]	46,XX,t(6;9)(p23;q34)[3]/46,sl,t(6;15)(p23;q21)[5]/47,sdl,+13[10]/46,XX[2]	+13[10/20]	No	No	Discordant

Note: Blank rows indicate longitudinal cases within the histories of the patient represented in the most recent occupied row.

Abbreviations: GPM: Genomic Proximity Mapping™, ISCN: International System for Human Cytogenomic Nomenclature, ELN: European Leukemia Network

Supplementary Table 6. Copy-neutral loss of heterozygosity (cnLOH) abnormalities detected by GPM compared with the corresponding clinical cytogenetic presentation. "Discordant" means the given abnormality was reported by clinical cytogenetics but missed or reported differently by GPM. "Added" means the given variant was reported by GPM but not by clinical cytogenetics.

Patient #	GPM findings	Clinical cytogenetics ISCN	cnLOH	ELN class-defining variant?	Other risk	Concordance between GPM and clinical cytogenetics
14	gpm 46,XY,dup(2)(p22.3p22.3)?c	46,XY[20].arr[GRCh38] 6pterp22.1(1_30,006,723)x2 hmz[0.9]	6p cnLOH	No	No	Discordant
9	gpm 46,XY,dup(1)(p36.22p36.22) ?c,inv(9)(p13.3p13.1)	46,XY[20].arr[GRCh38]13q12.13qter(27,055,669_ 114,338,054)x2 hmz[0.75]	13q cnLOH	No	No	Discordant
27	gpm 46,XX,del(2)(p23.3p23.3),del (16)(p13.11p12.3),(13)x2 hmz *	46,XX[20].arr[GRCh37] 2p23.3(24,587,652_26,417,829)x1, 13q12.11qter(19,814,912_115,103,529)x2 hmz, 16p13.13p12.3(12,040,511_18,539,704)x1	13q cnLOH	No	No	Concordant
35	gpm 47,XY,+8,(13)x2 hmz *	47,XY,+8[16]/46,XY[4].nuc ish(D8Z2x3)[161/200]	13q cnLOH	No	No	Added

* cnLOH only detectable by newer version software

Abbreviations: GPM: Genomic Proximity Mapping™, ISCN: International System for Human Cytogenomic Nomenclature, cnLOH: copy-neutral loss of heterozygosity, ELN: European Leukemia Network

Supplementary Table 7. Orthogonal evaluation of discordant copy number and structural variant calls by WGS testing, including whether WGS results support each sample's GPM results, clinically reported cytogenetics results, both, or neither. Constitutional abnormalities (indicated by "?c" per ISCN nomenclature) were not evaluated.

Patient #	GPM Findings (discordant calls bolded and underlined)	Clinical cytogenetics ISCN (discordant calls bolded and underlined)	Confirmation of presence/absence of discordant abnormality by Illumina WGS	Which platform's molecular presentation was corroborated by WGS?
2	gpm 46,XY,inv(16)(p13.11q22.1)	46,XY,inv(16)(p13.1q22)[13]/**92<4N>**,slx2[7].nuc ish(CBFBx2)(5'CBFB sep 3'CBFBx1)[119/200]/(CBFBx4)(5'CBFB sep 3'CBFBx2)[44/200]	Cannot confirm tetraploidy - cannot confirm GPM discrepancy	None
3	gpm 46,XY	47,XY,**+8**[15]/46,XY[5]	Low level trysomy 8 (roughly estimated <10%) present.	Clinical cytogenetics
5	gpm 46,XY	47,XY,**+8**[5]/45,X,-Y[3]/46,XY[12]	Unconfirmed +8	GPM
5	gpm 46,XY	47,XY,+8[5]/45,X,**-Y**[3]/46,XY[12]	Unconfirmed -Y	GPM
7	gpm 45,XY,**inv(3)(p24.3q26.2)**,-7[0.8]	45,XY,-7[11]/46,XY[9]	Confirmed inv(3)	GPM
11	gpm 46,XY,del(4)(q24q24)	47,XY,**+22**[2]/46,XY[18]	Unconfirmed trisomy 22	GPM
11	gpm 46,XY,**del(4)(q24q24)**	47,XY,+22[2]/46,XY[18]	Confirmed del(4)(q24q24)	GPM
16	gpm 46,XY,del(7)(q22.1q36.1),**dup(8)(p23.3)**,dup(12)(p13.31),inv(16)(p13.11q22.1) [0.8]	46,XY,del(7)(q22q36),inv(16)(p13.1q22)[17]/47,XY,inv(16)(p13.1q22),+22[3]	Confirmed by dup(8)(p23.3)	GPM
16	gpm 46,XY,del(7)(q22.1q36.1),dup(8)(p23.3),**dup(12)(p13.31)**,inv(16)(p13.11q22.1) [0.8]	46,XY,del(7)(q22q36),inv(16)(p13.1q22)[17]/47,XY,inv(16)(p13.1q22),+22[3]	Unconfirmed dup (12)	Clinical cytogenetics
16	gpm 46,XY,del(7)(q22.1q36.1),dup(8)(p23.3),dup(12)(p13.31),inv(16)(p13.11q22.1) [0.8]	46,XY,del(7)(q22q36),inv(16)(p13.1q22)[17]/47,XY,inv(16)(p13.1q22),**+22**[3]	Unconfirmed +22	GPM
18	gpm 46,XX,del(6)(p24.1p24.1),t(6;9)(p22.3;q34.12),t(6;15)(p24.2p24.1;q21.1),**d**	46,XX,t(6;9)(p23;q34)[3]/46,sl,t(6;15)(p23;q21)[5]/47,sdl,+13[10]/46,XX[2]	Unconfirmed dup 9	Clinical cytogenetics

	up(9)(p21.1p13.3)[0.3]			
18	gpm 46,XX,**del(6)(p24.1p24.1)**,t(6;9)(p22.3;q34.12),t(6;15)(p24.2p24.1;q21.1),dup(9)(p21.1p13.3)[0.3]	46,XX,t(6;9)(p23;q34)[3]/46,sl,t(6;15)(p23;q21)[5]/47,sdl,+13[10]/46,XX[2]	Confirmed del(6)(p24.1p24.1)	GPM
19	gpm 47,XY,+8	47,XY,+8[4]/46,sl,**-Y**[11]/46,XY[5]	Confirmed -Y	Clinical cytogenetics
24	gpm 46,XY	46,XY,t(2;19)(q35;p13.3)[5]/45,XY,**-21**[3]/46,XY[12]	Unconfirmed -21	GPM
24	gpm 46,XY	46,XY,**t(2;19)(q35;p13.3)**[5]/45,XY,-21[3]/46,XY[12]	Unconfirmed t(2;19)(q34;p13.3)	GPM
26	gpm 46,XX,t(10;17)(p11.2;q11.2)*[0.2]	46,XX,**add(17)(p13)**[2]/46,sl,del(10)(q24)[2]/46,XX,t(10;17)(p10;p10)[5]/46,XX[11].nuc ish(TP53,CEP17)x2[200]	Unconfirmed add(17)(p13)	GPM
26	gpm 46,XX,t(10;17)(p11.2;q11.2)*[0.2]	46,XX,add(17)(p13)[2]/46,sl,**del(10)(q24)**[2]/46,XX,t(10;17)(p10;p10)[5]/46,XX[11].nuc ish(TP53,CEP17)x2[200]	Unconfirmed del(10)(q24)	GPM
34	gpm 46,XX	46,XX,**del(7)(q31)**[5]/46,XX[15]	Unconfirmed del(7)(q31)	GPM
47	gpm 45,XY,der(3)del(3)(q21.3)inv(3)(q21.3q26.1)del(3)(q26.1q26.2),dup(4)(q32.2q32.2)?c,-7	45,XY,inv(3)(q21q26.2),-7[16]/45,sl,**del(6)(p23)**[4]	Unconfirmed del(6)(p23)	GPM

Abbreviations: GPM: Genomic Proximity Mapping™, ISCN: International System for Human Cytogenomic Nomenclature, WGS: whole genome sequencing